UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, STATION A, MONTRÉAL, QUÉBEC, H3C 3J7

TESTING THE DOGIT MODEL WITH AGGREGATE TIME-SERIES AND CROSS-SECTIONAL TRAVEL DATA

by

Marc J.I. Gaudry* and Michael J. Wills**

* Département de sciences économiques et Centre de recherche sur les transports, Université de Montréal.

** Strategic Planning Group, Transport Canada, Ottawa.

This research was partly supported by the Research and Development Centre of Transport Canada and by the F.C.A.C. program of the *Ministère de l'éducation du Québec*. Preliminary results were presented at the NBER-SNF Conference on Decision Rules and Uncertainty at Carnegie-Mellon University in March 1978.

Université de Montréal

Publication # 94 - Centre de recherche sur les transports
Cahier # 7801 - Département de sciences économiques

Mars 1978 - Révisé en août 1978

Dépôt légal - 2e trimestre 1978

ABSTRACT

This paper compares dogit and logit specifications of market share models, taking into account the possibility that conclusions might depend on transformations of the explanatory variables of these models. Parameter estimates are obtained both from a time-series urban transit mode of payment model and for a cross-sectional intercity mode choice model. It is demonstrated, using current maximum likelihood techniques extended to take multiple-order autocorrelation of the residuals into account, that the dogit specification is at least equal to, and sometimes clearly superior to, the logit specification irrespective of transformations of explanatory variables.

RÉSUMÉ

Cet article a pour but de comparer les modèles dogit et logit en tenant compte de transformations des variables explicatives pertinentes. Dans une première application, qui utilise les séries chronologiques, nous montrons que le modèle logit est inférieur au modèle dogit; dans la seconde, qui utilise une coupe instantanée, le modèle logit est presqu'aussi acceptable que le modèle dogit. Les transformations des variables ne modifient pas ces conclusions. Nos données agrégées décrivent la façon de payer les déplacements urbains par transport en commun et le choix entre modes de transport interurbains. Nous utilisons les méthodes courantes d'estimation du maximum de vraisemblance dont nous modifions la formulation pour tenir compte de processus autorégressifs des résidus d'ordre supérieur au premier.

CONTENTS

1. Introduction 1

2. Model specification. 1
   A. The dogit model 1
   B. The use of transformations. 2

3. Properties of the model. 4
   A. Admission of an "income effect" 4
   B. Reconciliation of compulsive and discretionary behavior 5

4. Selected applications. 7
   A. An urban transit mode of payment time-series model (MOP). 7
   B. An intercity mode of travel cross-sectional model (MOT) 8

5. Estimation procedure 8
   A. Rearrangement for estimation purposes and models estimated. . . . 8
   B. The likelihood function in the presence of multiple autocorrelation 10
   C. Maximization technique. 12

6. Results. 13
   A. Intended tests. 13
   B. Results for urban time-series MOP model 14
      B.1 General impact of dogit parameters. 14
      B.2 Comparison of nested model variants 14
      B.3 Conclusion for time-series model. 17
   C. Results for intercity cross-sectional MOT model 18
      C.1 General impact of dogit parameters. 18
      C.2 Comparison of nested model variants 18
      C.3 Conclusion for cross-sectional model. 21

7. Conclusion 21

Acknowledgements. 22

Bibliography. 23

Table 1. Results for urban time-series MOP model. 16

Table 2. Results for intercity cross-sectional MOT model. 20

Figure 1. Log likelihood surface for DCM variant of MOP model 15

Figure 2. Log likelihood surface for DCL variant of MOP model 15

Figure 3. Log likelihood surface for DU variant of MOT model. 19

Figure 4. Log likelihood surface for DCL variant of MOT model. 19

---

1. INTRODUCTION

This paper compares the ability of the logit and dogit models to explain observed aggregate market shares of travel alternatives. This comparison will take into account the possibility that results might be affected by transformations of the variables used as determinants of choices. We will show from time-series and cross-sectional data sets that the dogit model is at least as good as, and sometimes much better than, the logit model at explaining particular situations; we will also demonstrate that transformations of explanatory variables do not invalidate this result.

2. MODEL SPECIFICATION

A. The dogit model

Consider the following dogit market share model:

$$y_i = \frac{e^{V_i} + \theta_i \sum_j e^{V_j}}{(1+\sum_j \theta_j) \sum_j e^{V_j}} \quad , \quad i,j = 1,\ldots,M, \tag{1}$$

where y_i : share of the i^{th} of M alternatives;

θ_i : a non negative parameter associated with the i^{th} alternative: $\theta_i \geq 0$;

V_i : a function of K independent attributes X_{ik} associated with the i^{th} alternative.

This model, described in Gaudry and Dagenais (1977),does not require one to decide in advance whether the relative shares of two alternatives i and j are constant [@], as in the logit model: it is therefore in general

[@] In that sense it avoids or dodges the *a priori* decision of choosing a model format which commits one to a constant relative shares ratio or to a variable ratio — whence its name.

unconstrained by the independence from irrelevant alternatives axiom (IIA) which requires that the relative shares of a pair of alternatives depend solely on the characteristics of these alternatives. The ratio of market shares of alternatives determined by the dogit will be constant if the data are such that $\theta_i = \theta_j = 0$ for all alternatives — then the dogit collapses to the logit —, or if

$$\frac{\theta_i}{\theta_j} = \frac{e^{V_i}}{e^{V_j}} \quad , \quad \text{or} \quad \sum_j e^{V_j} = \sum_j e^{V_j^*}, \tag{2}$$

where the starred value represents a proposed change in the attributes of choices. Moreover, the dogit specification allows for constant relative shares of *particular* pairs of alternatives without simultaneously imposing this constancy on *all* relative shares. To establish the model, all that is necessary is to find that at least one parameter θ_i is significantly different from zero in any given choice problem.

B. The use of transformations

In specification (1), the functional form of the V_i functions is undetermined. We shall use the following:

$$V_i = \beta_{io} + \sum_k \beta_k \left[\frac{(X_{ik} + \mu_k)^{\lambda_k} - 1}{\lambda_k} \right] , \tag{3}$$

which can also be denoted as

$$V_i = \beta_{io} + \sum_k \beta_k \, X_{ik}^{(\lambda_k, \mu_k)} \tag{4}$$

where the transformation (λ_x, μ_x) of any variable X is defined as

$$X^{(\lambda_x, \mu_x)} = \begin{cases} \dfrac{(X+\mu_x)^{\lambda_x} - 1}{\lambda_x} & , \lambda_x \neq 0 \ , \ X + \mu_x > 0 \ , \\ \ell n \ (X + \mu_x) & , \lambda_x = 0 \ , \ X + \mu_x > 0 \ . \end{cases} \tag{5}$$

The μ_x is a location parameter studied by Tukey (1957) in the context of the power model. Box and Cox (1964) were among the first authors to apply the transformation as specified above. Schlesselman (1971) has emphasized the importance of including a constant term in (3) in order to preserve the invariance of the transformation to units of measurement. These transformations can be used to test the empirical validity of the theoretically derived functional form of a model or more simply to let the data themselves place restrictions on this form. Assume for instance that *a priori* considerations lead one to postulate the appropriateness of a logit model with a linear exponent; one might wish to test the linearity hypothesis of the model by specifying it as

$$y_i = \frac{e^{\beta_{io} + \sum_k \beta_k X_{ik}^{(\lambda_k, \mu_k)}}}{\sum_j e^{\beta_{jo} + \sum_k \beta_k X_{jk}^{(\lambda_k, \mu_k)}}} \quad . \tag{6}$$

Setting the vector $\mu = (\mu_1, \ldots, \mu_k, \ldots, \mu_K)$ equal to zero yields a Box-Cox logit form. If, in addition, the vector $\lambda = (\lambda_1, \ldots, \lambda_k, \ldots, \lambda_K) = 0$, the multiplicative or product model is obtained:

$$y_i = \frac{\exp(\beta_{io}) \prod_k X_{ik}^{\beta_k}}{\sum_j \exp(\beta_{jo}) \prod_k X_{jk}^{\beta_k}} \quad . \tag{7}$$

If $\lambda = (1, \ldots, 1, \ldots, 1)$ instead, the standard linear-in-parameters logit form is derived:

$$y_i = \frac{e^{\beta_{io} + \sum_k \beta_k X_{ik}}}{\sum_j e^{\beta_{jo} + \sum_k \beta_k X_{jk}}} \quad . \tag{8}$$

The authors have recently shown (Gaudry and Wills, 1977) elsewhere that format (8) gave the poorest explanation of the observed market shares of four intercity travel modes in Canada, the most convincing format being as in (6) and close to (7).

This paper will therefore test model (1) using the general formulation (3) of the V_i functions, thereby permitting an assessment of (1) within a general context. The set of parameters β_k, μ_k and λ_k associated with a particular explanatory variable X_k will not be made alternative specific because the numerous candidate models which result would obscure the central objective of the paper.

3. PROPERTIES OF THE MODEL

The specified model, obtained by substitution of (4) into (1), is

$$y_i = \frac{e^{\beta_{io} + \sum_k \beta_k X_{ik}^{(\lambda_k, \mu_k)}} + \theta_i \sum_j e^{\beta_{jo} + \sum_k \beta_k X_{jk}^{(\lambda_k, \mu_k)}}}{(1 + \sum_j \theta_j) \sum_j e^{\beta_{jo} + \sum_k \beta_k X_{jk}^{(\lambda_k, \mu_k)}}} \quad . \tag{9}$$

A number of its properties stated in the final version of Gaudry and Dagenais (1977) deserve restatement.

A. Admission of an "income effect"

The first property of the model can be summarized from the formulas for

own and cross elasticities of market shares, which are respectively

$$\eta_{ii} = \beta_k X_{ik} (X_{ik}+\mu_k)^{\lambda_k-1} \left[\frac{(1+\theta_i) e^{V_i}}{e^{V_i+\theta_i \sum_j e^{V_j}}} - \frac{e^{V_i}}{\sum_j e^{V_j}} \right] \qquad (10)$$

and

$$\eta_{ij} = \beta_k X_{jk} (X_{jk}+\mu_k)^{\lambda_k-1} \left[\frac{\theta_i e^{V_j}}{e^{V_i+\theta_i \sum_j e^{V_j}}} - \frac{e^{V_j}}{\sum_j e^{V_j}} \right] , \qquad (11)$$

where V_i and V_j are defined as in (3) and correspond to the exponents of (9). The θ_i parameters can be interpreted as indexes of the extent to which the attributes of all alternatives influence the choice of the i^{th} alternative. They function as an "income effect" which modifies the simplistic substitution effect built into the logit format in the following ways:

i) they reduce the η_{ii} and increase the η_{ij} which a logit format would have implied;

ii) this difference increases with the importance of the alternative the characteristics of which are altered.

B. Reconciliation of compulsive and discretionary behavior[@]

A second property can be seen by rewriting (9) as

$$y_i = \frac{\theta_i}{1+\sum_j \theta_j} + \frac{e^{V_i}}{(1+\sum_j \theta_j) \sum_j e^{V_j}} . \qquad (12)$$

[@]On November 16, 1977, Daniel Leblanc of L'Université de Montréal and Daniel McFadden of Yale University respectively suggested the analogy with Stone's expenditure system and the similar captivity interpretation.

This expression is compatible with the spirit of Stone's (1954) approach to demand analysis in which consumers first use a certain part of their income to acquire a set of goods at current prices, *whatever these may be*, and then distribute their remaining income over the set of available commodities in proportions which depend on prices. If $(\theta_i/(1+\sum_j \theta_j))$ is the share of income spent to satisfy "basic needs" for the i^{th} commodity, then $[1 - (\sum_i \theta_i/(1+\sum_j \theta_j))]$, or $[1/(1+\sum_j \theta_j)]$, is the share of discretionary income left over after the satisfaction of all basic needs; if a logit model further explains how discretionary income is spent, the sum of compulsive and discretionary elements determines the i^{th} commodity share of expenditure:

$$y_i = \frac{\theta_i}{1+\sum_j \theta_j} + \frac{1}{1+\sum_j \theta_j} \cdot \frac{e^{V_i}}{\sum_j e^{V_j}} \quad , \qquad (12')$$

which is clearly identical to (12).

This property of the dogit to distinguish between compulsive and discretionary behavior and to reconcile them in a single formula can also be understood as the property of allowing for a certain amount of "captivity" to alternatives. If a part of a travel market is "captive" to a mode and the rest of the market has a discretionary modal choice, the first term of (12) or (12') is interpreted as the "basic market" in the sense of the "basic need" above. This interpretation of the dogit has recently been formally investigated and extended by Ben-Akiva (1977) who also suggested that concepts of captivity have been used implicitly in a number of modal split studies. McFadden (1978) has recently interpreted the dogit as a special case of mixed probability random utility models.

One can therefore interpret the impact of the θ_i parameters on η_{ii}, the own elasticities of market share, as an income effect which decreases the impact of a substitution effect or as the captivity-induced reduction of the impact of alterations in characteristics: if θ_i were unbounded, η_{ii} would in the limit go to zero. The impact of θ_i on the cross elasticity of market share can be interpreted in similar fashion. The advantage of these interpretations is that they are relevant in binary applications, such as those chosen here, where the fact that the dogit is in general unconstrained by the IIA axiom is of little interest.

4. SELECTED APPLICATIONS

We will choose two distinct choice problems to test the dogit specification from aggregate data in two alternative contexts.

A. An urban transit mode of payment time-series model (MOP)

The first model explains the choice made by transit passengers between expensive tickets bought in small books,which require a low investment (e.g. 4 tickets for 1 dollar),and cheaper tickets sold in larger books,which require a larger investment (e.g. 12 tickets for 2 dollars). The probability of using either of these means of payment depends not only on price and required investment but also on the recurrence of activities for which trips are taken. These activities (work, school, shopping, etc.) exhibit seasonal fluctuations which we will capture with 12 monthly dummy variables. These constant seasonality factors are associated exclusively with the expensive mode of payment alternative. The sample consists of 80 consecutive monthly observations pertaining to the period May 1965 to December 1971 in Montreal.

B. An intercity mode of travel cross-sectional model (MOT)

The second model explains travellers' choice between the automobile and common carriers considered as one group (train, bus, plane) in terms of fare, travel time, departure frequency and an alternative-specific constant which will be associated with common carriers. The sample consists of observations on 56 Canadian city pairs for the year 1976. The distance between any two cities is never larger than approximately 1000 miles.

5. ESTIMATION PROCEDURE

A. Rearrangement for estimation purposes and models estimated

In the presence of two alternatives, taking logarithms of the ratio of the respective market shares y_1 and $y_2 = 1-y_1$ leads to the following form of (9):

$$\ell n \left[\frac{y_1 - \left(\theta_1 / 1+\theta_1+\theta_2\right)}{1-y_1 - \left(\theta_2 / 1+\theta_1+\theta_2\right)} \right]_t = V_t + u_t \quad , \tag{13}$$

where $V_t = (\beta_{10} - \beta_{20}) + \sum_k \beta_k \left((X_{1k_t} + \mu_k)^{(\lambda_k)} - (X_{2k_t} + \mu_k)^{(\lambda_k)} \right)$, (14)

t subscripts refer to observations 1, ..., N, and u_t is an error term.

We shall select for scrutiny a set of unconstrained and constrained dogit and logit models which are all variants of (13)-(14):

a) Dogit unconstrained or DU variant

In the unconstrained dogit model, the following region is admissible for the vector $\theta = (\theta_1, \theta_2) \geq 0$:

$$[\theta_1/1+\theta_1+\theta_2 < \text{min. } (y_1)] \text{ and } [\theta_2/1+\theta_1+\theta_2 < \text{min. } (1-y_1)]. \qquad (15)$$

In the MOP model, the elements of vector $\lambda = (\lambda_1, \ldots, \lambda_K)$ will be constrained equal but the elements of vector $\mu = (\mu_1, \ldots, \mu_K)$ will be left unrestricted except for the nonnegativity constraint $X_k + \mu_k > 0$ associated with the Box-Tukey transformation (5). In the MOT model, the reverse will be done.

a') Dogit equal unconstrained or DEU variant

As in (a) except that the elements of θ, i.e. θ_1 and θ_2 are constrained equal.

b) Dogit constrained linear or DCL variant

As in (a) except that the constraints $\lambda = 1$ and $\mu = 0$ are imposed. This corresponds to a linear exponent dogit specification.

c) Dogit constrained multiplicative of DCM variant

As in (a) except for the constraints $\mu = 0$ and $\lambda = 0$ which define the dogit version of product model (7).

d) Logit unconstrained of LU variant

As in (a) but imposing the constraint $\theta = 0$.

e) Logit constrained linear or LCL variant

As in (d) with the additional constraints $\mu = 0$ and $\lambda = 1$ which correspond to the standard linear exponent logit format.

f) Logit constrained multiplicative of LCM variant

As in (d) with the additional constraints $\mu = 0$ and $\lambda = 0$ which define the product model (7).

B. The likelihood function in the presence of multiple autocorrelation

If the error term vector u_t in (13) is assumed to be normally and identically distributed with a mean equal to zero and a constant variance σ^2, the likelihood of the original observations on market share y_1 can be formulated for fixed values of λ and μ as:

$$\mathcal{L}\ (y_1;\ \beta,\sigma^2,\theta) = (2\pi\sigma^2)^{-\frac{N}{2}} \cdot \exp\ (-\frac{u'u}{2\sigma^2}) \cdot |J\ (\theta;y_1)\ | \tag{16}$$

where $|\ J\ (\theta;y_1)|$ is the absolute value of the Jacobian of the transformation from the dependent variable in (13) to the actually observed y_1, or

$$|\ J\ (\theta;y_1)\ | = \prod_{t=1}^{N} \left\{(1+\theta_1+\theta_2)\ [y_{1_t} - \frac{\theta_1}{1+\theta_1+\theta_2}]\ [1-y_{1_t} - \frac{\theta_2}{1+\theta_1+\theta_2}]\right\}^{-1}$$

$$= \prod_{t=1}^{N} J_t\ . \tag{17}$$

It is convenient to maximize the log concentrated likelihood function

$$L(y_1;\theta) = -\frac{N}{2}\ [\ell n\ (2\pi)+1] - \frac{N}{2}\ \ell n\ [\frac{u'u}{N}] + \sum_{t=1}^{N} \ell n\ J_t \tag{18}$$

where

$$u_t = \ell n \left[\frac{y_1 - (\theta_1/1+\theta_1+\theta_2)}{1-y_1 - (\theta_2/1+\theta_1+\theta_2)}\right]_t - V_t\ , \tag{19}$$

which we can write

$$u_t\ =\ D(\theta)_t - \bar{V}(\lambda,\mu)_t\ . \tag{20}$$

This function is appropriate for the cross-sectional model but not for the time-series model for which a simplified likelihood function is derived.

This derivation will be outlined here but more details are found in Gaudry and Wills (1977) in the context of a Box-Cox or λ transformation of the dependent variable instead of a θ transformation as in this analysis.

If the error term u_t in (13) follows a stationary first order autoregressive process

$$u_t = \rho_1 u_{t-1} + e_t \ , \tag{21}$$

the concentrated log likelihood will be modified according to the well known format:

$$L(y_1; \theta, \rho_1) = -\frac{N}{2}[\ell n (2\pi)+1] - \frac{N}{2} \ell n \left[\frac{e'e}{N}\right] + \frac{1}{2} \ell n (1-\rho_1^2) + \sum_{t=1}^{N} \ell n J_t \ , \tag{22}$$

where $e'e = u'Q'Qu$ and the matrix Q is defined as

$$Q = \begin{bmatrix} \sqrt{1-\rho_1^2} & 0 & \cdots & 0 \\ -\rho_1 & 1 & \cdots & 0 \\ 0 & & \cdots & 0 \\ 0 & 0 & \cdots & -\rho_1 \ 1 \end{bmatrix} . \tag{23}$$

If the first observation from the sample is neglected, (22) becomes

$$L(y_1; \theta, \rho_1) = -\frac{N-1}{2}[\ell n (2\pi)+1] - \frac{N-1}{2} \ell n \left[\frac{e'e}{N-1}\right] + \sum_{t=2}^{N} \ell n J_t \ . \tag{24}$$

Beach and MacKinnon (1978) have recently pointed out that ignoring the first observation is less efficient in small samples than generally believed.

In the presence of higher orders of autocorrelation than the first, as with

$$u_t = \sum_{\ell=1}^{R} \rho_\ell \, u_{t-\ell} + e_t \, , \qquad \rho = (\rho_1, \ldots, \rho_\ell, \ldots, \rho_R) \, , \tag{25}$$

which constitutes a stationary AR(R) process in the sense of Box and Jenkins (1970), an extension of the simplification of the likelihood embodied in (24) yields

$$L\,(y_1;\, \theta, \rho) = -\frac{N-R}{2}\,[\ell n\,(2\pi)+1] \;-\; \frac{N-R}{2}\,\ell n\,\left[\frac{e'e}{N-R}\right] \;+\; \sum_{t=1+R}^{N} \ell n\, J_t \, . \tag{26}$$

This simplified likelihood avoids the computational difficulties associated with the incorporation of terms from the generalized autoregressive scheme in the full likelihood function at the cost of a certain loss of efficiency by dropping as many observations as R, the highest order of autocorrelation, requires.

C. Maximization technique

Likelihood functions (18) and (26) are maximized by nonlinear methods or more simply by scanning. A particular problem arises with the identification of the pertinent autoregressive scheme in the MOP time-series model which was left unspecified in (25). We identified the scheme (first and twelth orders) in the spirit of Box-Jenkins techniques by examining the residuals $\hat{u}_t$ obtained from applying least-squares to (19) for $\theta = \mu = 0$ and $\lambda = 1$. That specification determined (for fixed λ, θ and μ) a nonlinear equation obtained by combining (20) and (25):

$$D(\theta)_t - \sum_{\ell=1}^{R} D(\theta)_{t-\ell} = \hat{V}\,(\lambda,\mu)_t - \sum_{\ell=1}^{R} \hat{V}\,(\lambda,\mu)_{t-\ell} + e_t, \; t=1+R,\ldots,N, \tag{27}$$

which we solved by the Cochrane-Orcutt (1949) iterative procedure. Values of the likelihood function (26) were obtained by a grid search over values of θ, λ and μ without changing this initial specification of the autoregressive process. Expression (18) was maximized for the MOT model.

6. RESULTS

A. Intended tests

A natural test of the results as they pertain to the parameters θ, λ and μ is to compare the values of the maximized likelihood by means of the likelihood ratio test. Among the great number of possible conditional and unconditional tests on θ, λ, μ, β and ρ, we shall limit our explicit tests to the following, as summarized by their test statistics:

T.1 : DU vs LU

$$\hat{L}\ (\hat{\theta},\ \hat{\lambda},\ \hat{\mu},\ \hat{\beta},\ (\hat{\rho})) - \hat{L}\ (\theta=0,\ \hat{\lambda},\ \hat{\mu},\ \hat{\beta},\ (\hat{\rho}))\ , \tag{28}$$

T.2 : DCM vs LCM

$$\hat{L}\ (\hat{\theta},\ \lambda=0,\ \mu=0,\ \hat{\beta},\ (\hat{\rho})) - \hat{L}\ (\theta=0,\ \lambda=0,\ \mu=0,\ \hat{\beta},\ (\hat{\rho}))\ , \tag{29}$$

T.3 : DCL vs LCL

$$\hat{L}\ (\hat{\theta},\ \lambda=1,\ \mu=0,\ \hat{\beta},\ (\hat{\rho})) - \hat{L}\ (\theta=0,\ \lambda=1,\ \mu=0,\ \hat{\beta},\ (\hat{\rho}))\ , \tag{30}$$

T.4 : DEU vs LU

$$\hat{L}\ (\hat{\theta}_1=\hat{\theta}_2,\ \hat{\lambda},\ \hat{\mu},\ \hat{\beta},\ (\hat{\rho})) - \hat{L}\ (\theta=0,\ \hat{\lambda},\ \hat{\mu},\ \hat{\beta},\ (\hat{\rho}))\ . \tag{31}$$

T.1 compares unrestricted dogit and logit models whilst T.2 and T.3 pertain to the impact of λ and μ transformations on the relative merits of dogit and logit models; T.4 assesses the impact of an equality constraint on the dogit θ_i. The number of degrees of freedom associated with all of these tests is equal to 2 except for T.4 for which it is equal to 1. The statistic will be

$$[L(\Omega) - L\ (\omega)]\ >\ \frac{1}{2}\chi^2_\alpha\ (D.F.)\ , \tag{32}$$

where Ω is the parameter space under a maintained hypothesis, ω is the restricted space generated by a null hypothesis and $\chi^2_\alpha(D.F.)$ denotes the chi-square distribution with given degrees of freedom and significance level α. The percentage points of $X^2/2$ distribution with 2 degrees of freedom are

$$\begin{bmatrix} \text{Level} & 0.1 & 0.05 & 0.025 & 0.01 & 0.005 & 0.001 & 0.0001 \\ \text{Value} & 2.30 & 3.00 & 3.69 & 4.61 & 5.30 & 6.91 & 9.21 \end{bmatrix}\ . \tag{33}$$

B. Results for urban MOP time-series model

B.1 General impact of dogit parameters

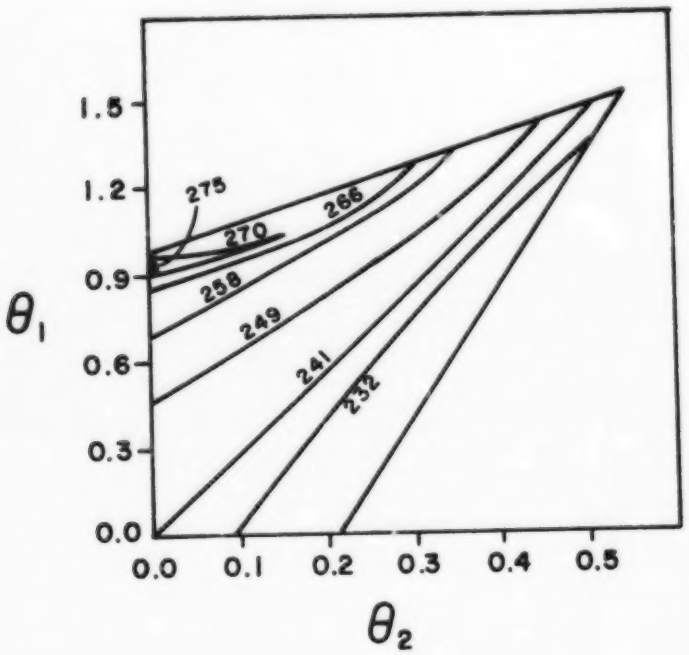
The general impact of dogit parameters $\theta = (\theta_1, \theta_2)$ on the urban transit MOP model is clearly shown in Figures 1 and 2 which exhibit the shape of the log likelihood function for various values of θ in variants DCM and DCL of the model. The surface is convex and asymmetric and shows the strong impact of dogit parameters as one approaches the constraints on admissible values of θ_1 and θ_2. Minimum values of high-price and low-price ticket shares y_1 and y_2 used in defining the feasible region for θ, according to (15), were respectively 50.11% and 17.48% in the sample. It is immediately apparent from Figures 1 and 2 that dogit results will be preferable to logit results because of the strong impact of the θ_i on values of the estimated log-likelihood.

B.2 Comparison of nested model variants

The results for various variants of model MOP are listed in Table 1. We note the following:

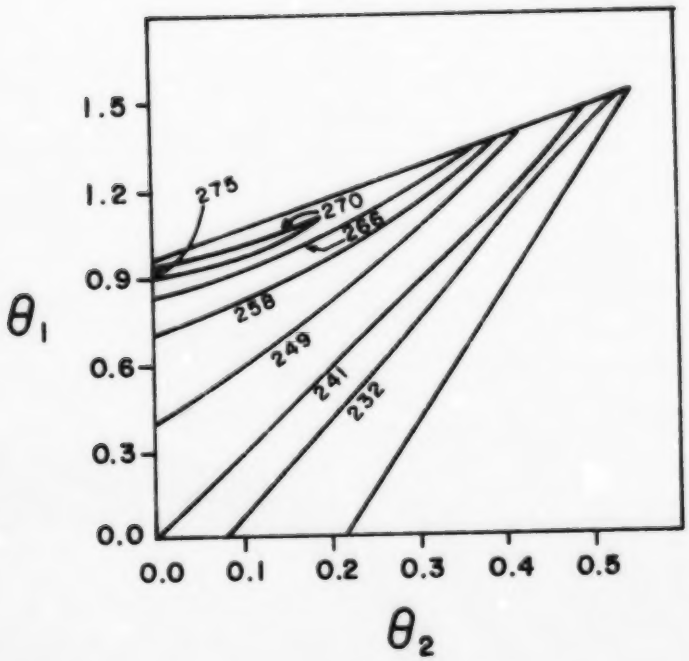
i) the value of the log likelihood function is so much higher for dogit variants than for logit variants that the logit format can be rejected with a very high degree of confidence. Indeed, the values of the test statistics for the pairwise comparison of unconstrained (DU), multiplicative (DCM) and linear exponent (DCL) dogit and corresponding logit forms (LU, LCM and LCL) are respectively: T.1 = 35.86, T.2 = 34.80 and T.3 = 34.63.

ii) the captivity effect associated with expensive low-investment ticket books is high (approximately 47%), which implies that the important working capital and risk of loss of quantity discounted tickets are



LOG LIKELIHOOD SURFACE FOR
DCM (θ, λ=0, μ=0) VARIANT OF MOP MODEL

FIGURE 1



LOG LIKELIHOOD SURFACE FOR
DCL (θ, λ=1, μ=0) VARIANT OF MOP MODEL

FIGURE 2

TABLE 1

RESULTS FOR URBAN TIME-SERIES MOP MODEL

PARAMETERS	MODEL VARIANTS*					
	DU	DCM	DCL	DEU/LU	LCM	LCL
θ_1 expens.tick.	.947	.943	.943	0.000	0.0	0.0
θ_2 cheap tick.	.000	.000	.000	0.000	0.0	0.0
λ_K price	-.052	0.0	1.0	.130	0.0	1.0
λ_K investment	-.052	0.0	1.0	.130	0.0	1.0
μ_K price	-.141	0.0	0.0	-.123	0.0	0.0
μ_K investment	-.670	0.0	0.0	-.899	0.0	0.0
price coeff.	-60.651	26.575	70.110	16.570	11.567	7.020
price t-stat.	(-5.16)	(3.18)	(1.60)	(7.07)	(6.71)	(9.38)
price elast.**	-10.68/11.16	0.78/-0.78	0.50/-0.49	10.68/-10.91	4.45/-4.45	6.73/-653
inv. coeff.	14.735	-3.94	-0.705	-0.127	1.203	0.391
inv. t-stat.	(5.32)	(-3.24)	(-2.26)	(-2.72)	(7.08)	(10.14)
inv. elast.**	3.76/-1.07	-0.11/0.11	-0.02/0.09	-0.98/0.08	0.46/-0.46	0.13/-0.69
12ct coeff.	42.50	-10.50	-6.60	-1.02	2.01	1.39
12ct t-stat.	(5.00)	(-4.50)	(-4.50)	(-2.90)	(6.60)	9.00
ρ_1 value	.931	.902	.896	.375	.328	.329
ρ_1 t-stat.	(17.09)	(14.30)	(13.80)	(3.40)	(2.93)	(2.95)
ρ_{12} value	.032	.056	.062	-.251	-.267	-.267
ρ_{12} t-stat.	(0.13)	(1.09)	(1.16)	(-2.32)	(-2.43)	(-2.43)
L (Ω)	276.95	275.62	275.44	241.09	240.82	240.81
L (Ω) - L (ω)	0.00	1.33	1.51	35.86	36.13	36.14
R^2 on D(ρ)****	.50	.47	.47	.75	.77	.77
R^2 on D****	.92	.95	.91	.80	.80	.80
D.F.	0	3	3	1/2	5	5

*The letters used in model names stand for: D = dogit, L = logit, C = constrained, U = unconstrained, M = multiplicative, L = linear exponent. These codes are more clearly defined in the text. t-statistics are computed in each variant for the given values of θ, λ and μ.

**The share own and cross elasticities of alternative 1 (expensive tickets) are given.

***In each model variant, coefficients and t-statistics of all 12 seasonal monthly factors were almost identical. True values differ from approximate mean values listed here by less than ± 10% for coefficients and less than ± 0.4 for t-statistics.

****Calculated on variables transformed by λ and μ in model format (13)-(14). R-squared should be used with caution as a measure of goodness of fit.

a strong barrier for the average transit user.

iii) the constrained equal dogit ($\theta_1 = \theta_2$) yields the same results as the logit, as shown in the DEU/LU column, which implies a test statistic equal to 0 for T.4 and emphasizes the importance of unrestricted estimation of θ. This is not surprising in a two-alternative situation: equal positive values of θ_1 and θ_2 would have implied, by Equation (2), equally attractive alternatives and constant equal market shares. It would not be a *trivial* result for a pair of alternatives in a multiple-choice situation.

iv) within logit and dogit subclasses, there are no statistical grounds for preferring one variant to another because the changes in the values of the log likelihood are extremely small.

v) from the point of view of the reasonableness of the results, the standard linear exponent logit LCL and the multiplicative logit LCM give the worst results because both explanatory variables (price and investment required) have positive signs. In all other variants at least one of these variables has the expected sign. This may reflect a specification flaw of the MOP model. The impact of the dogit format on the autoregressive scheme identified from the linear logit variant is considerable, as is its impact on share elasticities.

B.3 Conclusion for time-series model

The MOP model strongly supports the dogit specification and the logit specification can be rejected with an extremely high degree of confidence irrespective of the transformations assumed or estimated for the explanatory variables.

C. Results for intercity MOT cross-sectional model

C.1 General impact of dogit parameters

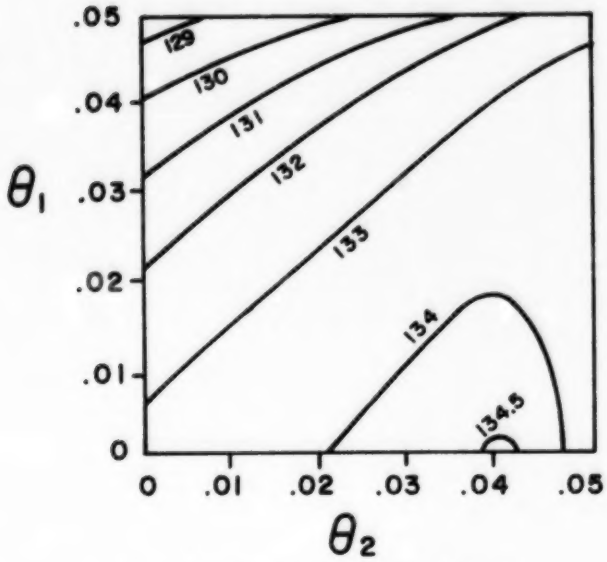
The general impact of dogit parameters $\theta = (\theta_1, \theta_2)$ on the intercity MOT model is clearly shown in Figures 3 and 4 which exhibit the shape of the concentrated log likelihood function for various values of θ in variants DU and DCL of the model. The results resemble those which were shown for the MOP model in that the surface is convex and asymmetric; they differ in the way in which the response occurs as one approaches the constraints on admissible values for θ. In interpreting these and subsequent results, one should keep in mind that, in (15), the minimum values of the automobile and common carriers' market shares in the sample were respectively $y_1 = 5.08\%$ and $y_2 = 5.22\%$. It is clear from Figures 3 and 4 that dogit results will not be very preferable to logit results.

C.2 Comparison of nested model variants

The results for various variants of model MOT are listed in Table 2. Apart from the fact that explanatory variables have expected signs in all model variants, we may note:

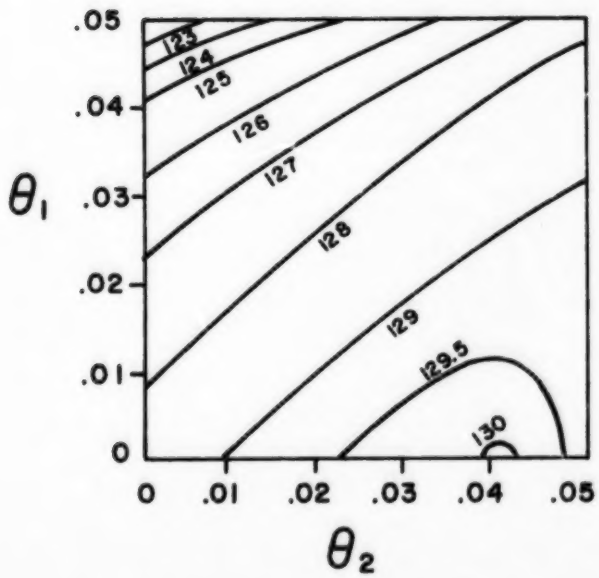
i) in its unconstrained (DU), multiplicative (DCM) and linear exponent (DCL) forms, the dogit model is only marginally preferable to the corresponding logit forms (LU, LCM and LCL): the values of the test statistics are respectively : T.1 = 1.21, T.2 = 2.09 and T.3 = 1.39, which implies that,with 2 degrees of freedom, the logit can only be rejected with a low degree of confidence.

ii) The dogit parameter θ_1 associated with the automobile is practically equal to zero but the parameter θ_2 is not, which implies that a weak captivity effect (4%) is associated with common carriers.



LOG LIKELIHOOD SURFACE FOR
DU (θ, $\lambda=\hat{\lambda}$, $\mu=\hat{\mu}$) VARIANT OF MOT MODEL

FIGURE 3



LOG LIKELIHOOD SURFACE FOR
DCL (θ, $\lambda=1$, $\mu=0$) VARIANT OF MOT MODEL

FIGURE 4

TABLE 2

RESULTS FOR INTERCITY CROSS-SECTIONAL MOT MODEL

PARAMETERS	MODELS VARIANTS*					
	DU	DCM	DCL	DEU/LU	LCM	LCL
θ_1 auto	10^{-5}	10^{-5}	10^{-7}	0.000	0.0	0.0
θ_2 public	.039	.044	.041	0.000	0.0	0.0
λ_k fare	.701	0.0	1.0	.786	0.0	1.0
λ_k time	.233	0.0	1.0	.362	0.0	1.0
λ_k frequency	-5.438	0.0	1.0	-4.674	0.0	1.0
μ_k all variables	52.08	0.0	0.0	41.44	0.0	0.0
fare coeff.	-2.394	-1.924	-1.530	-1.855	-1.545	-1.130
fare t-stat.	(-2.34)	(-2.90)	(-1.04)	(-2.06)	(-2.90)	(-0.95)
fare elast.**	-.77/.59	-.89/.89	-.42/.28	-.62/.45	-.77/.77	-.34/.23
time coeff.	-2.457	-1.093	-10.378	-2.960	-1.02	-9.485
time t-stat.	(-8.43)	(-4.67)	(-8.40)	(-9.79)	(-5.43)	(-9.44)
time elast.**	-.72/1.04	-.50/0.50	-1.09/3.71	-.80/1.32	-.51/.51	-1.08/3.67
freq. coeff.	17×10^{-12}	0.465	26 911	22×10^{-11}	0.371	21.603
freq. t-stat.	(2.91)	(2.55)	(1.07)	(2.97)	(2.53)	(1.05)
freq. elast.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.
ct auto coeff.	0.0	0.0	0.0	0.0	0.0	0.0
ct auto t-stat	(-)	(-)	(-)	(-)	(-)	(-)
ct pub. coeff.	-1.270	1.621	29.648	-1.136	1.273	23.743
ct pub. t-stat.	(-3.26)	(1.27)	(1.00)	(-3.64)	(1.24)	(0.98)
L (Ω)	134.52	131.78	129.95	133.31	129.69	128.56
L (Ω)-L(ω)	0.0	2.74	4.57	1.21	4.83	5.96
R^2(mean)***	.74	.70	.69	.77	.74	.70
R^2(origin)***	.82	.80	.79	.83	.81	.80
skewness****	2.45	3.10	2.59	2.09	2.56	2.06
kurtosis****	-0.28	0.15	-0.39	-0.18	0.24	-0.44
D.F.	0	4	4	1/2	6	6

*The letters used in model names stand for: D = dogit, L = logit, C = constrained, U = unconstrained, M = multiplicative, L = linear exponent. These codes are more clearly defined in the text. t-statistics are computed in each variant for the given values of θ, λ and μ.

**The share own and cross elasticities for alternative 2 (public) are given for market shares assumed to be equal.

***Calculated on variables transformed by λ and μ in model format (13)-(14).

**** Skewness and kurtosis for the distribution of errors.

iii) Constraining the elements of θ to be equal yields the same results as the logit in the DEU/LU column, which implies a test statistic equal to 0 for T.4 and emphasizes the importance of unrestricted estimation of θ.

iv) Within dogit and logit subclasses, there are weak grounds for preferring common multiplicative or linear exponent formats to unrestricted formats because the unrestricted forms DU and LU remove four restrictions from common specifications DCM-LCM and DLC-LCL.

C.3 Conclusion for cross-sectional model

In constrast with what was found for the MOP model, the MOT model does not support the dogit specification with a high degree of confidence and this conclusion is not changed by removing functional form constraints typically used in logit model exponents. Mode specific constants may partly explain the relatively good performance of the logit model because such constants in effect reduce the level of captivity to be explained by dogit parameters: in that sense, MOT results might have been anticipated.

7. CONCLUSION

We have shown in the time-series MOP model that the logit specification should be rejected in favor of the dogit specification. The latter specification was only marginally preferable to the former in the cross-sectional MOT model. These results may indicate that a minimum observed market share of 5% for all alternatives may be a rough upper bound for the convincing use of the logit model the "tails" of which are often too "thin" in particular applications. The distribution of observed market share sizes is of

importance in this context: a very low minimum and a very high maximum of the share of one of the alternatives will, by (15), limit the range of θ; a doubly-peaked distribution would suggest partitioning the sample and using the dogit in the "thick-tail" part and the logit elsewhere, but such a choice takes us beyond the scope of this paper. It should be noted that the admissible region (15) for the θ_i depends on individual observations and that, if these contain error, the estimate of the θ_i will be biased. One solution to this problem is to drop suspect observations, keeping in mind the decisive role of the smallest market shares in the sample. We have also found that transformations of explanatory variables did not alter our conclusions and provided only weak grounds of choice among logit and dogit variants.

ACKNOWLEDGEMENTS

The views presented in this paper are the authors' and in no way engage the institutions with which they are affiliated. The authors are jointly responsible for the whole of the analysis but Marc Gaudry, of the *Université de Montréal* assumes primary responsibility for the time-series application and Michael Wills, of the Strategic Planning Group of Transport Canada, Ottawa, for the cross-sectional application. Marcel G. Dagenais provided valuable advice. Danielle Bilodeau made a significant contribution as research assistant and Richard Laliberté performed computations for the time-series model. The research on which this article is based was supported partly by the Research and Development Centre of Transport Canada and by the F.C.A.C. program of the *Ministère de l'éducation du Québec*. Partial results were presented at the NBER-NSF Conference on Decision Rules and Uncertainty at Carnegie-Mellon University in March 1978.

BIBLIOGRAPHY

Beach, C.M. and MacKinnon, J.G. (1978). A maximum likelihood procedure for regression with autocorrelated errors. Econometrica 46, 51-58.

Ben-Akiva, M. (1977). Choice models with simple choice set generating processes. Working paper. Department of civil engineering, M.I.T.

Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. Journal of the Royal Statistical Society, Series B, 211-243.

Box, G.E.P. and Jenkins, G.M. (1970). Time-Series Analysis, Forecasting and Control, Holden-Day, San Francisco.

Cochrane, D. and Orcutt, G.H. (1949). Application of least-squares regression to relationships containing autocorrelated error terms. Journal of the American Statistical Association 44, 32-61.

Gaudry, M.J.I. and Dagenais, M.G. (1977). The dogit model. Publication # 82, Centre de recherche sur les transports, Université de Montréal. Forthcoming in Transportation Research.

Gaudry, M.J.I. and Wills, M.J. (1977). Estimating the functional form of travel demand models. Publication # 63, Centre de recherche sur les transports, Université de Montréal. Forthcoming in Transportation Research.

McFadden, D. (1978). Econometric models for probabilistic choice. Paper presented at the NBER-NSF Conference on Decision Rules and Uncertainty, Carnegie Mellon University. Mimeographed. p. 19.

Schlesselman, J. (1971). Power families: a note on the Box and Cox transformation. Journal of the Royal Statistical Society, Series B., 33, 307-311.

Stone R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of British demand. The Economic Journal 64, 511-527.

Tukey, J.W. (1957). On the comparative anatomy of transformations. Annals of Mathematical Statistics 28, 602-632.